Universitat de Lleida

Escola Politècnica Superior

Enginyeria Informàtica

Treball de Final de Carrera

Votació electrònica amb recompte homomòrfic

Víctor Mateu Meseguer

Lleida - Maig de 2010

Tutors

JOSEP M. MIRET BIOSCA

FRANCESC SEBÉ FEIXAS

CAPíTOL 1

Introducció

La societat humana és una societat en constant evolució. Cada cop ens separem més del comportament basat en la subsistència i ens apropem més a un comportament passiu basat en el respecte. En aquesta evolució també s'han vist modificats alguns patrons de conducta social molt importants. Antigament les societats eren molt piramidals, a la punta d'aquesta piràmide trobàvem una família amb poder que, normalment, era l'encarregada de prendre les decisions que afectaven a la resta de la "seva" societat. A mesura que la societat s'anava fent més gran, la societat anava guanyant poder i la presa de decisions es complicava cada vegada més. Per superar aquesta dificultat es creaven nous estats socials que s'encarregaven de facilitar la feina de la familia "dominant". Amb el temps i la creació de les grans societats, aquest sistema va resultar no ser prou efectiu, i va evolucionar a una primera idea de democràcia, que vindria a ser la possibilitat per al poble d'escollir els membres, o ideologies, que prendran les decisions. El poble no podia prendre les decisions, pero sí que podia decidir qui volia que s'encarregués de fer-ho. Amb la intenció de facilitar aquesta elecció neix el concepte de votació. La idea principal d'aquesta és donar el mateix valor a l'opinió de qualsevol persona vinculada amb la decisió a prendre. Quan ens trobem davant d'una decisió que s'ha de prendre en grup, acostumem a votar per prendre la decisió que beneficiï a més membres del grup, aquest és un exemple senzill de votació.

Quan la votació no té més trascendència, la transparència d'aquesta pot fer-la més senzilla i per tant resulta útil. No obstant en votacions importants hi ha alguns requisits que eviten que la votació pugui ser transparent. En unes votacions per escollir alcalde, per exemple, les persones que aspiraven a ser escollides per entrar al poder podrien

optar per coaccionar a un votant per obligar-lo a votar el que ells volen en comptes de guanyar-se la confiança d'aquest. Com el vot d'aquest votant seria públic tothom podria comprovar què ha votat, deixant al votant desprotegit.

El que es vol és protegir als votants, fer que el seu vot estigui completament desvinculat d'ells, però assegurant que el seu vot s'ha tingut en compte. Amb aquest objectiu trobem les votacions actuals, en les que el vot va tancat dins d'un sobre i es diposita dins de l'urna amb la resta dels vots. Així aconseguim l'anonimat del votant i amb l'ajuda del cens garantim que el votant formi part de la votació i no voti més d'un cop.

Actualment, el sistema aconsegueix els objectius de confidencialitat i robustesa, però encara hi ha problemes a superar. Un d'aquests problemes, potser el més evident, és la necessitat de centralitzar el lloc de votació, és a dir, tothom s'ha de desplaçar fins el lloc on es dugui a terme la votació per exercir el seu dret a votar. És incòmode però, en principi, tothom pot fer l'esforç.

Relacionat amb aquest primer entrebanc ens trobem amb el segon, la coacció. Hem aconseguit que ningú pugui saber quin vot hi ha dins del sobre del votant, malgrat tot, un coaccionador pot omplir un sobre amb el vot i donar-lo al votant coaccionat mentre comprova que efectua el vot correctament. Si tothom va a votar al mateix lloc, el coaccionador pot obligar a alguns votants a votar el que ell els demani, ja que òbviament no li costa res identificar les seves víctimes. Si això succeeix, tenim un grup de votants que no només veu privat el seu dret a votar, també veu com el seu vot reforça una opció de vot que no comparteix. Aquesta dificultat encara existeix en l'actualitat i per aquest motiu la llei persegueix als coaccionadors.

Tenint en compte els entrebancs de la votació i amb totes les mesures utilitzades per a que aquesta es dugui a terme correctament, ens trobem amb el pas posterior a la votació, el recompte. El cost en temps de fer el recompte de vots és significatiu i representa una dificultat que intentem reduir al màxim. El recompte el fan un conjunt de persones, en molts casos els mateixos que verifiquen el procés de votació, i òbviament inverteixen

molt de temps en fer el recompte. Com més persones menys temps es triga, però com més vots més costa fer el recompte, de manera que es busca un equilibri entre persones recomptant en funció del tamany del cens que acaba amb un elevat nombre de persones invertint moltes hores en el procés complet de votació. A més, un error en el recompte pot passar inadvertit, tot i que hi ha persones encarregades de vetllar per a que no passi.

Òbviament es poden trobar altres problemes quan l'organització de la votació no realitza correctament la seva feina. En aquests casos ens podem trobar amb cens incorrectes que provoquen irregularitats en la votació. No obstant aquests problemes no es donen gairebé mai.

Amb la votació en la situació actual, i amb la intenció d'aprofitar els avantatges que ofereixen les noves tecnologies dins del món de la votació, neix la votació electrònica. La idea és fer exactament els mateixos passos, però de manera segura, remota i automatitzada, aprofitant els beneficis de la tecnologia.

El primer problema que afrontem en aquest camí per instaurar la votació electrònica és la confiança dels votants. La societat ha vist com les màquines poden ser manipulades i també té present que aquestes, alguns cops, no funcionen com es desitja. Consideren que hi ha votacions on el que es vota és massa important i per tant no és tolerant a errors.

Per solventar aquest problema s'han estat realitzant estudis per trobar-hi solucions. Aquí apareix la criptografia, que permet assegurar al votant que ningú podrà saber el seu vot, garantint-li la confidencialitat del vot. Malgrat tot, un criptosistema millor no els dóna totes les garanties que ells necessiten. Una votació no es pot simplificar només al vot d'una persona, el que interessa també és que la votació no estigui manipulada i és en aquest punt on neix la desconfiança.

Per garantir que la votació no ha estat manipulada s'estan dissenyant diferents sistemes que ofereixen garanties en la correctesa dels resultats de la votació. En el nostre projecte

treballem amb dos d'aquests sistemes: la votació amb mixing i prova de correctesa del mixing, i la votació amb recompte homomòrfic i validació de vots.

1. Objectius i treball realitzat

Hem volgut dissenyar i implementar un sistema de votació amb recompte homomòrfic, un dels sistemes més rápids de votació que garanteix les propietats de seguretat que ha de mantindre una votació electrònica. A més hem implementat una prova de seguretat en els vots dels emisors per a garantir que el recompte s'ha fet correctament.

Per dur a terme el projecte s'han realitzat dues tasques en paral·lel:

- Aprenentatge:
  - Estudiar la xifra de Paillier.
  - Estudiar votacions amb recompte homomòrfic.
  - Analitzar problemes relacionats amb aquests recomptes.
  - Estudiar proves de validació de vots amb coneixement nul.
  - Analitzar el rendiment de les proves que demostren que un missatge pertany a un conjunt (MLS, "Message Lies in Set").
- Implementació:
  - Implementar un aplicatiu que simuli una votació electrònica.
  - Adaptar l'aplicatiu per a que pugui treballar amb qualsevol criptosistema.
  - Implementar el recompte homomòrfic
  - Implementar la prova MLS sobre els vots rebuts.

L'objectiu és comparar el rendiment d'una votació segura utilitzant mixing amb els d'una votació segura utilitzant recompte homomòrfic.

La memòria conté tres capítols d'explicació i un de conclusions i treball futur. En el primer fem un breu repàs a l'evolució de la votació electrònica, mentre que en els dos següents parlem dels tipus de votació amb mixing i amb recompte homomòrfic, explicant els conceptes bàsics que necessitarem per entendre les diferències dels resultats i el significat d'aquests.

CAPíTOL 2

Votació electrònica

La tecnologia ha donat la possibilitat d'incrementar les capacitats de la votació tradicional en tots els sentits. Aquestes millores són tan evidents que han fet proliferar molt ràpidament els interessos en la democràcia electrònica. Això provoca que es comencin a realitzar estudis per la seva legalització i s'accelera les inversions en investigació en aquest camp. Però la votació electrònica no es limita al context de les eleccions de govern, també resulta molt útil en les votacions que es duen a terme en organitzacions, tant públiques com privades, que es troben distribuïdes geogràficament.

Entrant a parlar de votació electrònica, parlarem de les fases en què dividim el procés de votació, cadascuna amb uns procesos interns que pretenen simular els processos d'una votació tradicional. Aquestes fases són:

- Fase de preparació.

  (1) Anunci de l'elecció: L'autoritat de l'elecció publica el propòsit de l'elecció, la llista de candidats, requisits, condicions, etc.

  (2) Registre de votants: Es busca aconseguir informació sobre els votants per a la seva posterior identificació.

- Fase de votació.

  (1) Autenticació del votant: El votant s'autentica per poder emetre el vot.

  (2) Establiment de sessió: Es relaciona el votant amb el servidor de vot al que es conecta per efectuar la votació.

(3) Selecció de candidats: Es mostra la llista de candidats per a que el votant esculli.

(4) Enviament del vot: Un cop escollit el candidat, es xifra l'elecció del votant i s'envia al servidor corresponent per a la seva tramitació.

(5) Validació del vot: Es verifica que el votant només ha votat un cop i que el vot emés és correcte.

- Recompte i publicació

(1) Transferència de vots: Quan ha acabat la fase de votació l'autoritat de votació transfereix els vots a l'autoritat d'escrutini.

(2) Permutació de vots: S'utilitza una funció de permutació de vots per trencar qualsevol lligam entre el votant i el seu vot. S'acostuma a combinar amb un rexifratge.

(3) Desxifratge de vots: Es recupera el vot original per al seu posterior recompte.

(4) Escrutini.

(5) Auditoria de resultats: Comprovació que els resultats són correctes i que no hi ha hagut cap manipulació tant en la recol·lecció com en el recompte.

(6) Publicació de resultats.

Hi ha diferents tipus de votació electrònica. Amb el temps se n'han anat creant de nous a mesura que les noves tecnologies i els coneixements de seguretat ho han permés.

1. Tipus

En primer lloc parlarem, per antiguitat, del sistema de votació basat en paper però amb recompte automàtic. Aquest va ser la primera evolució del sistema de vot normal en què la gent porta el seu vot dins d'un sobre, que serà dipositat en una urna després

de comprovar que el votant es troba dins del cens. El que pretenia aquest sistema de votació era facilitar el compte de vots fent-lo automatitzat. La idea era utilitzar sistemes de recompte de vots mitjançant l'escaneig òptic i/o tabulació electrònica, de manera que els votants podien emplenar la seva butlleta manualment i després la màquina en faria el recompte ràpidament.

Més endavant la tecnologia va evolucionar fins al punt de permetre als votants generar les seves pròpies butlletes amb una màquina. Aquest sistema de votació, tot i ser força efectiu, tampoc representava una evolució remarcable en el món de la votació ja que l'únic problema que resol és el del recompte però segueix forçant a les persones a desplaçar-se.

Un altre tipus de votació és la votació electrònica de registre directe. En aquest cas el votant es troba amb un dispositiu, semblant a un petit ordenador, que li permetrà autenticar-se i realitzar el seu vot directament, sense butlleta ni urna. La idea és fer més accesible el vot a totes les persones. Un cop enregistrats tots el vots, aquests queden guardats dins de la memòria del dispositiu, des d'on es poden enviar a qualsevol sistema de centralització de vots. Normalment el dispositiu compta els vots a mesura que els votants els van inserint, de manera que quan els ha d'enviar, els envia ja comptats i deixa una còpia en memòria per una possible revisió. Els dispositius han anat evolucionant a mesura que s'han trobat defectes en el sistema, i avui per avui ja disposen de sistemes de seguretat tant en la transferencia de vots, com en la recol·lecció, per evitar manipulacions no desitjades en els recomptes. En aquests sistemes, la necessitat de desplaçar-se per part del votant no s'ha solucionat.

Una evolució de l'anterior sistema és la votació electrònica de registre directe utilitzant una xarxa pública. Ara els dispositius envien la informació a través d'Internet. Serà important que les dades es transmetin de forma segura i per tant, s'afegiran mecanismes de xifratge i signatura digital als dispositius per assegurar que la votació ha estat feta correctament. Aquest tipus de votació ja s'ha utilitzat durant força temps en les

votacions internes d'algunes empreses i també en votacions en departaments d'àmbit públic.

Per últim, tenim la votació electrònica per xarxa pública, en la que els votants disposen de maquinari per poder fer la votació utilitzant Internet. Sense cap dubte és la més complicada en quant a seguretat, però també la més favorable al votant ja que no li cal desplaçar-se. És en aquest camp en el què aprofundirem més. El repte és aconseguir una votació des de qualsevol lloc amb un nivell de seguretat encara més elevat que en el cas de la votació tradicional. S'ha de tenir en compte que s'haurà d'utilitzar un software específic per dur a terme aquestes votacions i per tant caldrà assegurar-se que no només els nostres sistemes informàtics són segurs, sinó que el votant ho està fent sense cap possibilitat de ser "espiat".

2. Requisits de seguretat

La principal sospita que recau sobre un procés electrònic és la seva seguretat en front d'un procés que es realitza manualment. Per tant qualsevol possibilitat d'intervenció no desitjada en el procés ha de ser controlada, tant si prové de fora com de dintre del propi sistema.

Els següents punts constitueixen una descripció del requisits de seguretat que ha de complir una votació electrònica:

Legitimitat del votant: Només poden participar votants autoritzats i només es tindrà en compte un vot per votant. Normalment s'utilitza un cens per identificar els votants autoritzats. En el cas de la votació electrònica s'utilitzen tècniques d'identificació remota.

Robustesa: Un votant només pot realitzar un vot i ningú fora del cens pot efectuar una votació.

Confidencialitat: La relació entre el votant i el seu vot no pot ser coneguda ni deduïble.

Precisió: El resultat de la votació ha de venir donat pel recompte de tots els vots recollits de manera legítima. S'ha d'evitar qualsevol alteració dels vots mitjançant sistemes de prevenció i detecció d'alteracions.

No informació: No es poden conèixer resultats parcials durant la votació ja que el coneixement de l'estat de la votació podria influir en els votants que encara no han votat.

Verificació: Pot ser de dos tipus:

- Individual: El votant ha de poder comprovar que el seu vot ha estat enviat, rebut i processat correctament.
- Universal: És important que hi hagi una comprovació pública per a que qualsevol participant i/o observador pugui verificar la integritat del resultat.

Evitar coercions: Un votant no hauria de poder demostrar a un tercer el vot que ha escollit, ja que una tercera persona podria exigir al votant que votés una opció en concret. Si no ho pot demostrar aquesta tercera persona no pot saber si ha votat el que ell ha exigit o no.

CAPíTOL 3

Criptografia

Des dels temps de l'Antic Egipte fins avui dia, ha passat molt de temps, però hi ha coses que no han canviat, com és el desig de l'ésser humà d'amagar els seus secrets.

Personatges com Cleopatra o César van aprendre la importància d'amagar els seus missatges de les mirades indiscretes. La Scitala que els espartans utilitzaven prop del 400 a.C, o el propi codi César van ser els principis. El 1466 León Battista Alberti va idear el sistema polialfabètic basat en la rotació d'uns "corrons". Un segle més tard Giovani Battista Belaso va inventar la clau criptogràfica basada en una paraula o text que es trancrivia lletra a lletra sobre el missatge original.

El naixement de la informàtica i dels criptosistemes informàtics va suposar un canvi radical del concepte de criptografia, i també del criptoanàlisi. Els criptosistemes i els algoritmes van augmentar considerablement la seva complexitat. Des del DES fins als criptosistemes asimètrics basats en corbes el·líptiques hi ha hagut molts canvis.

Criptografia, l'art d'amagar, té el seu origen en el grec: kryptos (amagat) i graphein (escriure). L'art d'amagar un missatge mitjançant signes convencionals és molt antic, gairebé tant com l'escriptura. Claude E. Shannon va publicar en dos anys, dos documents que van suposar la fundació de la Teoria de la Informació [**16**], [**17**].

La criptografia es complementa amb una altra branca d'estudi, el criptoanàlisi, que estudia el camí invers de la criptografia. Dedica els seus esforços a desenmascarar els secrets que la criptografia intenta amagar.

En el context criptogràfic, considerem com a text en clar qualsevol informació que resulta llegible i comprensible. Un text en clar seria qualsevol informació abans de ser encriptada

o després de ser desencriptada. Es considera que qualsevol informació és vulnerable si es troba en aquest estat. En aquest mateix context, considerem com a criptograma, qualsevol informació que es trobi convenientment xifrada, i no resulti llegible ni comprensible més que per al destinatari legítim de la mateixa.

Al mecanisme per transformar un text en clar en un criptograma l'anomenen xifratge. De la mateixa manera anomenem desxifratge al procés de recuperar la informació a partir d'un criptograma.

En un criptosistema la informació segueix sempre un mateix flux.

(1) L'emissor xifra el text en clar, i envia el criptograma resultant.

(2) El criptograma és transmés per un canal insegur fins arribar al receptor.

(3) El criptograma arriba al receptor, que el desxifra i obté el text en clar.

Hi ha diferents tipus de criptosistemes:

Algoritmes simètrics: Són els criptosistemes més senzills. Es tracta d'algoritmes que treballen amb una única clau amb doble funció. Dins d'aquests criptosistemes podem distingir entre dos tipus d'algoritmes: els de xifratge de bloc, i els de xifratge en flux. Tot i no parlar de la implementació dels algoritmes, sí que citarem els principals algoritmes criptogràfics simètrics:

- DES [**18**] (Data Encryption Standard): Algoritme de 64 bits de clau, dels quals 56 componen la clau del xifratge, mentre que els 8 restants són de paritat i s'utilitzen per detectar errors. Actualment DES ja no és estàndar criptogràfic. Va ser trencat al gener de 1999, amb un sistema de còmput que analitzava 250.000.000.000 claus per segon.

- Triple DES [**1**]: Degut a la capacitat de còmput actual i la relativa facilitat que suposa trencar el DES, es desenvolupa un sistema de triple aplicació de l'algoritme DES, amb 3 claus diferents per aplicar succesivament. De

fet, s'utilitza una clau externa dividida, ja que DES matemàticament no és un grup, i la seva aplicació repetida provoca un increment efectiu del tamany. Amb aquest sistema s'obté un xifratge de 192 bits, 168 efectius i 24 de paritat.

- AES (Rijndael) [**9**] (Advanced Encryption Standard): Es tracta d'un algoritme simètric de xifratge de blocs de longitud variable. Se serveix de claus de longitud variable: 128, 192 o 256 bits. El 26 de Novembre de 2001 i amb un procés d'estandardització que va durar 5 anys va ser anunciat pel NIST com FIPS 197 dels Estats Units. Es va fer estàndard efectiu el 26 de Maig de 2002.

- IDEA [**3**] (International Data Encryption Algorithm): Va ser creat al 1990 per X. Lai i L.Massey. Es tracta d'un algoritme simètric de xifratge en blocs de 64 bits. El seu funcionament es basa en operacions senzilles com multiplicacions d'enters, sumes i XOR. Treball amb claus de 128 bits.

- RC4: Va ser dissenyat per Ron Rivest de la RSA Security l'any 1987. RC4 forma part dels protocols de xifratge més comuns com són WEP, WPA per a targes wireless i TLS. Els principals avantatges que ofereix són simplicitat a l'hora d'implementar-lo i molta velocitat en la seva execució. Es tracta d'un algoritme de xifratge en flux.

Funcions Hash: A més dels criptosistemes, podem parlar d'algoritmes criptogràfics, la finalitat dels quals acostuma a ser més de verificació que de xifratge i desxifratge. Dins d'aquest concepte podem incloure els hash, que tenen unes característiques que els fan especialment útils per a tasques de verificació. Les propietats més importants d'un hash són:

  - Unidireccional: Conegut un hash, cadena de bits extrets a partir d'una altra cadena de bits, és computacionalment impossible la reconstrucció del missatge original.

- Compressió: A partir d'un missatge de qualsevol longitud s'obté un hash de tamany fix.
- Difussió: El resum és una funció complexa de tots els bits del missatge.
- Col·lissió simple: Donat un missatge qualsevol, és computacionalment impossible trobar un altre missatge, el resum del qual sigui el mateix.
- Col·lissió forta: És computacionalment dificil trobar dos missatges amb un resum idèntic.

Alguns dels principals algoritmes criptogràfics de resum són:

- MD5 [**14**] (Message Digest 5): Va ser ideat pel matemàtic Ron Rivest el 1992, i suposa l'evolució dels algoritmes MD2 i MD4. Es tracta d'una funció criptogràfica de tipus hash que accepta com a entrada un missatge de qualsevol tamany i retorna com a sortida una cadena de 128 bits.
- SHA-1 [**10**] (Secure Hash Algorithm - 1): Va ser ideat pel NIST el 1994 com una aplicació de l'algoritme SHA. Es tracta d'una funció criptogràfica de tipus hash que accepta una entrada de 2^{64} bits com a màxim, i retorna una cadena de 160 bits. És lleugerament més lent que MD5 però també és computacionalment més complex de trencar i per tant més segur.

Algoritmes asimètrics: Són criptosistemes més moderns i complexos que els simètrics, i per tant més segurs. Es basa en l'existència d'un parell de claus complementàries de manera que un criptograma xifrat per una de les claus només pot ser desxifrat per l'altra. Els algoritmes asimètrics més importants són:

- RSA [**15**] Algoritme de xifratge de clau pública desenvolupat al 1977. A l'actualitat és l'algoritme asimètric més utilitzat i és vàlid tant per xifrar com per firmar digitalment. La seva seguretat es basa en la dificultat de factoritzar nombres enters.

- ElGamal [**2**] És un algoritme de xifratge de clau pública. La seguretat de l'algoritme es basa en la suposició que calcular un logaritme discret té una complexitat computacional molt alta. El procediment de xifratge/desxifratge està basat en càlculs sobre qualsevol grup cíclic finit G. Això fa que la seguretat del procediment depengui directament de la dificultat de calcular un logaritme discret a G.

- Paillier [**11**] Va ser presentat el 1999 per Pascal Paillier. Algoritme de xifratge de clau pública que hem emprat per a la realització del projecte. La seva seguretat es basa en la dificultat computacional de trobar un enèsim residu. L'introduim més profundament en el següent capítol.

Els dos criptosistemes anteriors tenen unes propietats que els fan molt útils quan treballem en votacions electròniques, ja que dos missatges (vots) iguals poden tindre un missatge xifrat diferent. Aquesta propietat resulta especialment important per a mantindre la confidencialitat dels vots. A més tenen operacions homomòrfiques que utilitzades correctament ens seran molt útils per optimitzar els recomptes i fer enmascaraments.

1. La xifra de Paillier

El sistema criptogràfic de Paillier és un algoritme probabilístic (que basa el seu resultat en alguns components aleatoris) utilitzat en la criptografia de clau pública. La seva seguretat es basa en que el problema computacional de trobar l'enèsim residu és computacionalment difícil.

L'esquema és un criptosistema homomòrfic additiu, més endavant parlarem de les propietats que ens ofereix.

Per a més informació sobre el criptosistema, consultar [**11**].

1.1. Inicialització.

Per inicialitzar el sistema necessitem generar les claus, tant la pública com la privada. Per fer-ho seguim els següents passos:

(1) Escollim 2 nombres primers grans, aleatoris i independents l'un de l'altre, tals que compleixin que $mcd(p \cdot q, (p-1), (q-1)) = 1$.

(2) Calculem $N = p \cdot q$.

(3) Calculem $\lambda = mcm(p-1, q-1)$.

(4) Escollim un enter aleatori $g \in \mathbb{Z}^*_{N^2}$

(5) Ens assegurem que N divideix l'ordre de g comprovant l'existència de la següent multiplicació modular inversa:

$$\mu = (L(g^{\lambda} \pmod{N^2}))^{-1} \pmod{N},$$

on la funció L ve definida per:

$$L(u) = \frac{u-1}{N}.$$

Un cop fet això, podem donar la clau pública (N, g) mentre que guardarem la clau privada (λ, μ) en secret.

1.2. Xifratge.

En aquest mateix context tenim m que és el missatge a xifrar, on $m \in \mathbb{Z}_N$. Per xifrar el missatge seguirem els següents passos:

(1) Seleccionem una r aleatòria tal que $r \in \mathbb{Z}^*_N$.

(2) Calculem el missatge xifrat c mitjançant la funció $c(m, r) = g^m \cdot r^N \pmod{N^2}$.

Com es pot veure, la quantitat de missatges xifrats diferents que poden correspondre a un mateix missatge és $\phi(n)$. Per tant tot i que tinguem un mateix missatge xifrat diversos cops, com és el cas d'una votació, un atacant no podria relacionar els missatges xifrats entre sí.

1.3. Desxifratge.

Donat un missatge xifrat c, tal que $c \in \mathbb{Z}^*_{N^2}$. Obtenim el missatge desxifrat m a partir de la funció:

- $m = L(c^\lambda \pmod{N^2}) \cdot \mu \pmod{N}$.

Fixem-nos en que per desxifrar no es té en compte el valor de r, que era la component aleatòria del nostre text xifrat. Aquesta propietat ens ajudarà a emmascarar els nostres vots.

1.4. Emmascarament.

Considerem l'emmascarament, com la propietat d'un criptosistema per aconseguir, a partir d'un missatge xifrat c, un altre missatge xifrat c' tal que al desxifrar c i c' el missatge obtingut sigui el mateix. Per aconseguir-ho aprofitarem el fet que treballem amb un criptosistema additiu homomòrfic on es cumpleixen les propietats següents:

- Donat un missatge xifrat

$$c(m, r) = g^m \cdot r^N \pmod{N^2}.$$

  Si $m = 0$ llavors $c(0, r) = r^N \pmod{N^2}$.

- Donat un missatge xifrat $c(m, r) = g^m \cdot r^N \pmod{N^2}$ i un altre missatge xifrat $c'(m', r') = g^{m'} \cdot r'^N \pmod{N^2}$ podem calcular el missatge xifrat $c'' = c \cdot c'$. Al

resultat de desxifrar c'' li direm m'' tal que

$$m'' = L((c \cdot c')^{\lambda} \pmod{N^2}) \cdot \mu \pmod{N}$$

i es compleix la propietat que $m'' = m + m'$.

La idea és obtenir un $c'' \neq c$ tal que $m'' = m$. Per fer-ho necessitem un c' on $m' = 0$. Per tant $c'' = c \cdot c' = g^m \cdot r^N \cdot r'^N \pmod{N^2}$.

Així doncs per emmascarar un missatge xifrat c el que farem serà:

(1) Seleccionem una r aleatòria tal que $r \in \mathbb{Z}_N^*$.

(2) Generem el missatge emmascarat c' com $c' = c \cdot r^N \pmod{N^2}$.

1.5. Homomorfisme aditiu. Parlem d'homomorfisme aditiu quan operant els text xifrats ens resulta una suma dels texts en clar. En el cas del Paillier:

- La múltiplicació de dos texts xifrats es desxifrarà com la suma dels texts en clar:

$$D(E(m, r) * E(m', r') \pmod{N^2}) = m + m' \pmod{N}.$$

- Un text xifrat elevat a la potencia d'un text en clar es desxifrarà com la multiplicació dels texts en clar:

$$D(E(m, r)^{m'} \pmod{N^2}) = m \cdot m' \pmod{N}.$$

Aprofitant aquestes dues propietats podem realitzar votacions amb recompte homomòrfic, com veurem en els següens capítols.

CAPíTOL 4

Tècniques de recompte

Com s'ha parlat abans, una votació consta de diferents fases, des de la publicació del cens fins a la verificació dels resultats. També hem parlat de les condicions de seguretat que ha de tenir una votació electrònica. Sabem que la utilització dels criptosistemes ens ajudarà a amagar l'elecció del votant, no obstant aquesta primera capa no serà suficient per garantir l'anonimat dels nostres votants. Si nosaltres rebem un vot xifrat i després el desxifrem, existirà un vincle directe entre el codi del vot xifrat i el seu contingut, que juntament amb la relació entre votant i el seu vot xifrat ens acabarà dient què ha votat cadascú. En l'actualitat hi ha dues formes d'afrontar aquesta situació, els mixings i els recomptes homomòrfics.

1. Mixings

Les votacions amb mixing (mescla de vots) proposen la idea conceptualment més pròxima a les votacions tradicionals. Podem parlar d'un paral·lelisme conceptual entre els diferents instruments de la votació. En aquest cas la butlleta és el missatge a xifrar; el xifratge seria el sobre tancat i l'estructura de dades que enmagatzema els vots seria l'urna. En una votació tradicional, les urnes s'agiten per evitar que els encarregats de fer el recompte puguin saber el vot d'algun votant, el sobre del qual han estat vigilant dins de l'urna. En el cas d'alguns criptosistemes el vot xifrat té una identitat única i per tant resulta més senzill encara vincular-lo amb el votant. El que es coneix com a mixing és el procés en el qual els vots xifrats es desordenen i s'enmascaren obtenint un nou conjunt de vots xifrats que no tenen cap relació deduïble amb els vots enviats pels votants.

El cost de fer el mixing és despreciable ja que barrejar un conjunt de vots és una operació amb un cost molt baix i que creix de manera lineal, mentre que fer el rexifratge de cada vot té un cost més elevat pero també creix de manera lineal en funció del nombre de vots i del tamany de la clau del criptosistema i per tant parlaríem d'un cost assumible.

Tenim el problema de l'anonimat del votant solucionat, però acabem de generar-ne un de nou. Què passa si quan barrejem els vots i els reencriptem, el que fem realment és generar v vots nous i canviar-los pels vots existents? Com podem garantir que això no passa? La solució són les proves de correctesa.

1.1. Proves de correctesa. La prova de correctesa s'utilitza per demostrar que els vots no s'han modificat durant el mixing. Hi ha diferents tipus de proves de correctesa: les proves interactives i les no interactives. Les primeres requereixen un intercanvi d'informació entre el verificador i el provador que en entorns de votació limitarien les persones que podrien realitzar aquestes proves. Les no interactives, per contra, el que fan és generar variables noves a partir de les dades "secretes" de la votació i oferir les operacions necessàries per a que amb les dades públiques de la votació i les variables generades puguin verificar que el mixing s'ha realitzat correctament. Aquestes proves són especialment útils en votació ja que la prova que publica l'entitat encarregada de realitzar el mixing pot ser verificada per qualsevol persona o entitat.

Durant el projecte hem utilitzat la implementació realitzada en [**5**] de la prova de correctesa presentada a [**12**]. Es tracta d'una prova amb un cost que té un creixement quadràtic, i que prova que la mescla s'ha realitzat correctament, que no s'ha modificat cap vot i permet saber quin vot s'ha modificat en cas que el provador detecti alguna irregularitat.

2. Recomptes homomòrfics

Els recomptes homomòrfics constitueixen la segona opció per desvincular la identitat del votant amb el contingut del vot xifrat. En l'apartat anterior hem vist com el mixing

conseguia resoldre el problema però ens obligava a realitzar una prova de correctesa que influia en el cost de la votació. En el cas dels recomptes homomòrfics el que fem és utilitzar l'operació homomòrfica del criptosistema per generar un únic text xifrat que és el resultat d'aplicar aquesta operació en tots els vots xifrats.

A l'apartat 3.1 de la memòria hem vist com funcionava l'operació homomòrfica amb la xifra de Paillier i com utilitzem aquesta propietat per enmascarar vots. La idea principal era multiplicar dos text xifrats $c(m, r)$ i $c'(m', r')$ on $m' = 0$ de manera que obteníem $c''(m + m', r'')$ tal que al desxifrar c'' el resultat que obtenim és $m + m' = m$.

Si els valors del missatges m_i possibles els definim en funció del nombre de votants aconseguim que la propietat homomòrfica del sistema ens permeti fer el recompte directament. Un exemple en una votació molt petita: Tenim una votació amb 4 candidatures (A,B,C,D) i un cens de 10 persones. Implementem una votació en la que per votar a A, cal prendre $m = 1$, per votar a B, $m = 10$, per votar a C, $m = 10^2$, i per votar a D, $m = 10^3$.

Suposem que:

- $m_0 = 1$, $m_1 = 10$, $m_2 = 10$, $m_3 = 10$, $m_4 = 10^2$, $m_5 = 10^2$, $m_6 = 10$, $m_7 = 10$, $m_8 = 10^2$ i $m_9 = 10^3$.
- el votants envien $c_i(m_i, r_i)$.

Operem homomòrficament tots els vots i obtenim:

$$c_{TOTAL}(\sum_{i=0}^{9} m_i, r_{TOTAL})$$

Desxifrem c_{TOTAL} i obtenim el resultat com la suma única dels missatges de tots els votants.

En aquest cas l'anonimat i la privadesa del votant s'assegura sempre, ja que el seu vot no és desxifrat mai i per tant resulta impossible associar el seu vot xifrat amb el seu

vot en clar. També evitem el problema de confiar en una entitat de votació que pugui intentar canviar vots, ja que tothom pot calcular c_{TOTAL}. L'últim dels avantatges que ens ofereix és que el recompte es realitza molt ràpidament ja que el servidor només ha de fer un únic desxifratge i fer el recompte a partir de divisions.

El problema que tenim en aquest cas, i que es deriva del recompte homomòrfic, recau en la validesa dels vots. En l'exemple anterior tothom ha enviat un c_i que xifrava un missatge m_i vàlid, però que passaria si per exemple $m_4 = 4 \cdot 10^2 - 3 \cdot 10$? En aquest cas el cinquè votant hauria votat 4 vegades C, i no només això, també anul·laria a tres persones que haurien votat la candidatura B, i el pitjor de tot és que ningú se n'hauria pogut adonar i per tant la votació es consideraria correcta.

Afortunadament aquest problema té solució però el seu cost, tot i crèixer de manera lineal, no és del tot eficient. El que hem de fer és que cada votant provi que el seu vot pertany al conjunt de vots vàlids (proves MLS).

2.1. Proves MLS. Ens referim amb aquest nom a l'algoritme que permet demostrar que el text en clar del criptograma forma part d'un conjunt de texts en clar possibles. Com hem vist anteriorment, quan treballem amb votacions amb recompte homomòrfic, els missatges que podem xifrar han de presentar una forma concreta. Si un votant maliciós utilitza un missatge fora del conjunt de missatges "vàlids" pot manipular el resultat de la votació. En aquest context neix la necessitat de poder demostrar que el text en clar que conté un criptograma pertany a un conjunt.

Les proves MLS es poden realitzar sobre diferents criptosistemes, en el nostre cas treballem amb la xifra de Paillier ja que és la que ens permet aprofitar millor les propietats homomòrfiques del recompte.

Per implementar la prova MLS sobre Paillier utilitzem un algoritme on hi ha dos personatges que intercanvien informació durant la prova: el provador i el verificador. El

primer ha de convencer al segon de que el text xifrat,

$$c = g^{m_i} \cdot r^N \pmod{N^2}$$

encripta un missatge que pertany al conjunt $S = \{m_1, m_2, ..., m_p\}$.

- El provador escull una $\rho \in \mathbb{Z}_N^*$. També escull aleatoriament p - 1 valors $\{e_j\}_{j \neq i} \in \mathbb{Z}_N$ i p - 1 valors $\{v_j\}_{j \neq i} \in \mathbb{Z}_N^*$.
- El provador calcula $u_i = \rho^N \bmod N^2$ i $\{u_j = v_j^N (g^{m_j}/c)^{e_j} \bmod N^2\}_{j \neq i}$.
- El provador genera $e_{chall} = H(u_1, ...u_p)$ i calcula:
  - $e_i = e_{chall} - \sum_{j \neq i} e_j \bmod N$.
  - $v_i = \rho \cdot r^{e_i} \cdot g^{(e_{chall} - \sum_{j \neq i} e_j) \div N} \bmod N^2$, on $\div$ és el quocient de la divisió entera.
- El provador envia $\{u_j, v_j, e_j\}_{j \in \{1,..,p\}}$ i e_{chall} al verificador.
- El verificador comprova que $e_{chall} = \sum_j e_j \bmod N$.
- El verificador comprova que $v_j^N = u_j (c/g^{m_j})^{e_j} \bmod N^2$ per cada $j \in \{1, ..., p\}$

CAPíTOL 5

Detalls d'implementació

En l'anterior projecte de votació electrònica [**5**] es va implementar un sistema capaç d'admetre diferents criptosistemes i proves per realitzar una votació. Aprofitant que el codi s'havia dissenyat amb la idea de ser escalable, el primer que hem fet ha estat ampliar els criptosistemes per a permetre proves MLS.

En primer lloc hem implementat les funcionalitats de generar i verificar una prova d'aquestes característiques. En el nostre cas, com treballem sobre Paillier, hem realitzat la prova MLS sobre aquest criptosistema. Després hem fet la part de verificació per a tindre completa la funcionalitat de la classe.

Recordem que teníem una classe ColegiElectoral que agafava vots i en feia la mescla. Ara el que necessitem és que no faci aquesta mescla, sino que verifiqui la prova i faci el recompte homomòrfic. Pensant en l'escalabilitat, dissenyem un nova jerarquia de "colegis", amb una interfície principal "Colegi" que pugui rebre vots i fer el recompte posterior. A partir d'aquesta interficie, la classe ColegiElectoral que ja teníem passa a implementar aquesta interfície i creem una nova classe ColegiMLS que també implementa la interfície Colegi. En el nostre cas, i seguint els criteris per desvincular classes, el que fem és que en el recompte es realitzi la prova MLS de manera que el resultat sigui transparent a l'usuari.

Pel que fa a la jerarquia de vots, només hem afegit una nova clase VotPaillierMLS que implementa la interfície VotEncriptat que ja havíem creat en el projecte anterior. A aquesta nova clase hi hem afegit les provesMLS de manera que el col·legi corresponent pugui utilitzar-les quan sigui necessari.

Durant el procés d'adaptació hem trobat alguns petits detalls del disseny anterior que eren millorables, com la creació d'una jerarquia de votants per treballar amb diferents tipus de vots. Com hem parlat en l'apartat 4.2, els valors dels vots per fer un recompte homomòrfic són molt concrets i per tant el votant hauria d'enviar el seu vot amb el valor correcte. Per evitar errors de votació el que fem és crear una jerarquia de votants que envia vots, de manera que s'assegura que el vot està "ben format".

El programa realitzat estructura la seva execució en tres parts:

Inicialització de variables: On s'escull el criptosistema i el sistema de votació que utilitzarem.

Recollida de vots: La clase Votant genera vots aleatoris ben formats i els envia a la classe Colegi.

Recompte: En aquest punt es poden donar dues situacions:

- Si la votació és amb mixing els vots es mesclen i es recompten normalment, mostrant els resultats de la votació.
- Si es una votació amb recompte homomòrfic, el que es fa primer és verificar la prova MLS de cada vot per assegurar la validesa de la votació, i després fem la multiplicació i desxifrem el resultat per trobar el recompte final de la votació.

Verificació: En les votacions amb mixing, després de publicar els resultats, es prova que el mixing s'ha realitzat correctament utilitzant la prova de coneixement nul implementada a [**5**].

CAPíTOL 6

Comparativa de resultats

Volem veure quin sistema resulta millor en cada situació. Per comprovar-ho necessitem fer les proves en funció de 3 variables diferents:

- En primer lloc tindrem en compte la influència del tamany de la clau en l'eficiència del sistema. Com més gran és la clau, més seguretat donem al criptosistema i més complicats es fan els càlculs de les proves.
- En segon lloc veiem com influeix el nombre de votants. Si una votació perd eficiència amb el nombre de votants haurem de saber afitar la votació per tal d'escollir la més eficient.
- Per últim analitzem el cost d'afegir possibilitats de vot a la votació. En aquest cas només ens interesa veure quan el recompte homomòrfic deixa de millorar els temps de la votació amb Mixing.

Dividim els resultats en dues gràfiques. La primera mostra l'eficiència dels sistemes quan treballen amb claus de 512 bits. La segona mostra els mateixos resultats però sobre una seguretat de 1024 bits en la clau del Paillier.

Veiem com la votació amb mixing és més eficient quan el nombre de vots possibles supera els 10. La votació homomòrfica presenta una eficiència molt bona amb un nombre de vots reduïts, però es mostra poc tolerant a l'increment de vots possibles ja que quan augmentem el nombre de possibilitats del votant el temps del recompte creix molt ràpidament. Tot i això el creixement de l'homomòrfic en funció dels votants és més reduït i per tant a partir del miler de votants ja resulta més eficient per a 10 possibles candidats al vot.

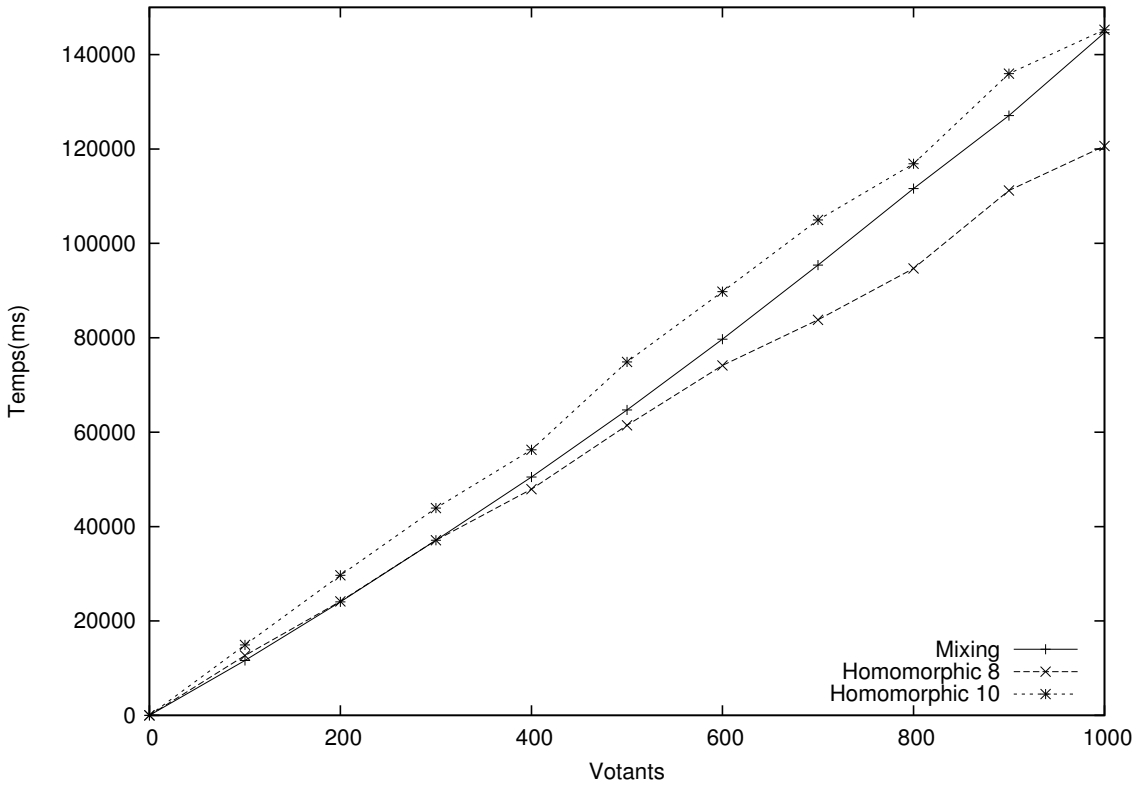

FIGURA 1. Comparativa amb clau de 512 bits

Hem realitzat les mateixes proves incrementant i reduint el nivell de seguretat del criptosistema per veure l'incidència que té en els temps dels dos sistemes de votació. Ens trobaríem amb una situació interessant si a l'incrementar la seguretat ens permetés augmentar el nombre de vots possibles. No obstant això no és el que passa, com veiem a la Figura 2. El que ens trobem és que el nombre de vots possibles s'ha de reduir per sota de vuit si volem mantenir la rapidesa del mixing en la votació amb recompte homomòrfic. Com més incrementem la seguretat més reduït serà el nombre possibilitats i per tant la votació es mostra intolerant als increments en el nivell de seguretat del criptosistema.

En aquesta segona gràfica també mostrem els resultats d'una votació amb només 3 possibles candidats. Veiem com funciona el sistema de recompte homomòrfic dins d'un

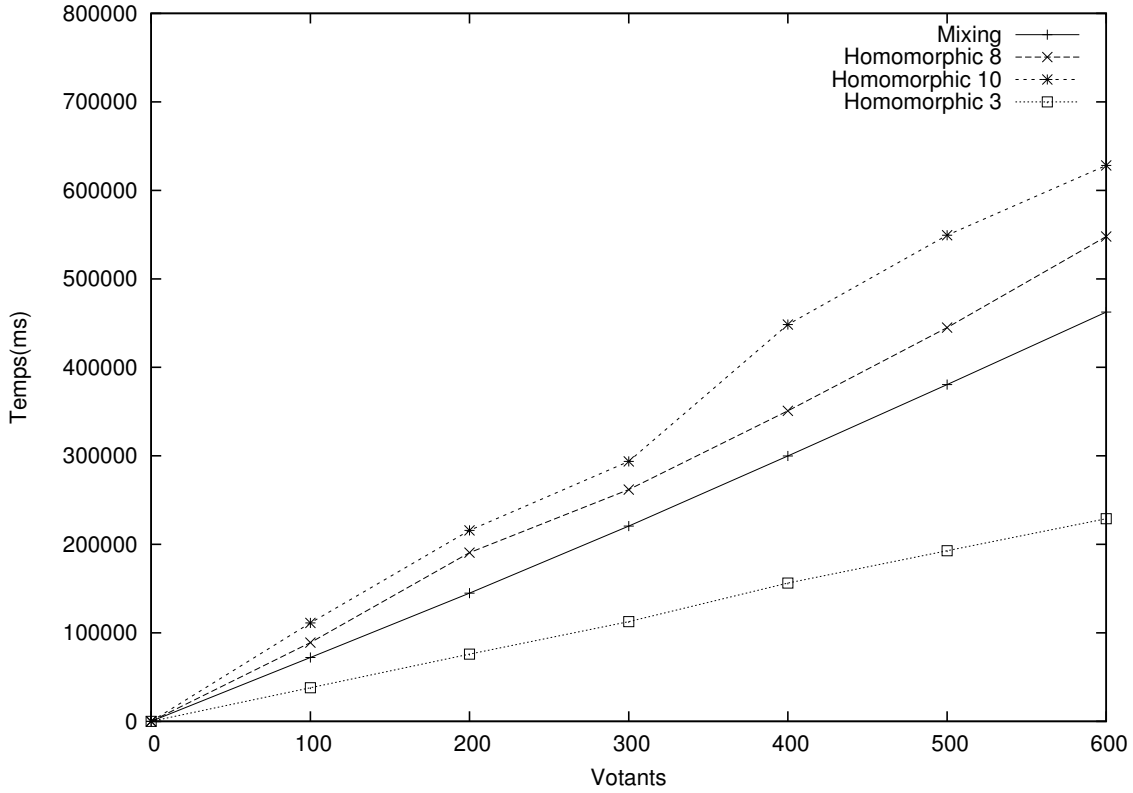

FIGURA 2. Comparativa amb clau de 1024 bits

tipus de votació real que afavoreixi al màxim l'eficiencia de l'algorisme. En tipus concrets de votació, el recompte homomòrfic ens ofereix una millora de rendiment important respecte les votacions amb mixing.

Els resultats, en general, mostren com el recompte homomòrfic és més ràpid que el mixing però les proves necessàries per a cada sistema tenen una eficiencia que fa que la prova MLS resulti menys eficient quan es supera un llindar de possibles missatges. Podem veure com la prova MLS té una dependència directa amb el nombre de candidats, mentre que la prova de correctesa d'un mixing només depèn del volum de votants de la votació.

CAPíTOL 7

Conclusions

Després de veure els resultats obtinguts a l'hora de realitzar les proves, la conclusió més evident és que la prova de correctesa que presenta [**12**] utilitzant la xifra de Paillier depèn de menys factors que la prova MLS implementada per al recompte homomòrfic.

Trobarem exemples de votació en les que el recompte homomòrfic ofereix un rendiment superior, però també trobarem altres casos en els que el cost es dispara al comparar-lo amb el del mixing. La part positiva és que els avantatges de la prova MLS són els mateixos que els de la votació amb recompte homomòrfic i per tant en les situacions en que el recompte homomòrfic sigui un element de votació a tindre en compte, la seva prova serà també la més eficient. Parlaríem de sinergia entre el sistema de recompte i la seva prova.

Podem afirmar que els sistemes de votació amb mixing i prova de correctesa són més volubles i permeten una varietat de missatges molt gran a més d'una prova robusta que ens permet fer un seguiment més proper a l'evolució de cada vot, però sense comprometre la seguretat de la votació. D'altra banda les votacions amb recompte homomòrfic i prova MLS ens ofereixen un rendiment molt millor quan treballem amb un nombre de candidats afitat i no molt nombrós. Òbviament aquesta rapidesa no compromet la seguretat de la votació i l'auditoria resulta molt més senzilla, tant que qualsevol persona pot auditar una votació amb recompte homomòrfic.

1. Treball futur

Seguint la línia del treball que s'ha realitzat, podríem ampliar la recerca en diferents direccions:

- Millorar el rendiment d'una votació amb recompte homomòrfic intentant implementar una verficació de la prova MLS en batch per reduir el temps de la prova en el servidor.
- Implementar un sistema híbrid i comprovar l'eficiència d'aquest envers els dos sistemes per separat.
- Comparar el rendiment d'un sistema amb mixing i una prova de correctesa més lleugera que utilitzi els homomorfismes per demostrar la correctesa del mixing.
- Utilitzar la implementació de tot l'aplicatiu per adaptar-lo a l'entorn web i utlitzar-lo per a fer diferents tipus de votació depenent de la tipologia d'aquesta.
- Implementar servidors que realitzin la verificació de la prova MLS i que utilitzin signatura cega per distribuir la part amb més carrega de les votacions homomòrfiques.

Bibliografia

[1] American National Standards Institute, *Triple Data Encryption Algorithm Modes of Operation* ANSI X9.52-1998.

[2] T. ElGamal, *A public key cryptosystem and a signature scheme based on discrete logarithms.* IEEE trans. Inform. Theory, 31, no. 4 , 469-472, 1985.

[3] X. Lai, *On the design and security of block ciphers*, ETH Series in Information Processing, J.L. Massey (editor), vol. 1, Hartung-Gorre Verlag Konstanz, Technische Hochschule (Zurich), 1992.

[4] D. Master, *Criptosistemas informáticos*, 2004.

[5] V. Mateu, *Implementació d'un sistema de votació sobre la xifra de Paillier i ElGamal.* Treball de final de màster, UdL, 2009.

[6] A. J. Menezes, P.C. V. Oorschot, S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press New York, 1997.

[7] V. Morales, *Seguridad en los procesos de voto electrónico remoto* PhD. Thesis UPC - Departamento de Telemàtica, 2009.

[8] R. Moreno, J. Pujolàs, P. Sanz, M. Serio. *Mix verificables con pares ElGamal y curvas elípticas.* Actas VI Jornadas de Matemática Discreta y Algorítmica, 469-476, 2008.

[9] National Institute of Standards and Technology. FIPS Pub 197: *Advanced Encryption Standard (AES).* November 2001.

[10] National Institute of Science and Technology, Federal Information Processing Standard (FIPS) *Secure Hash Standard*, 180-1, April 1993.

[11] P. Paillier *Public-key Criptosystems Based on Composite Degree Residuosity Classes* Procs of EUROCRYPT'99 , LNCS vol.1592, pp. 223-238, 1999.

[12] K. Peng, C. Boyd, E. Dawson, *Simple and Efficient Shuffling with Provable Correctness and ZK Privacy* Procs of CRYPTO'05 , LNCS vol.3621, pp. 188-204, 2005.

[13] A. Riera, *Design of Implementable Solutions for Large Scale Electronic Voting Schemes* PhD. Thesis UAB, 1999.

[14] R. Rivest, *The MD5 Message-Digest Algorithm*, RFC1321, MIT LCS and RSA Data Security, Inc., April 1992.

[15] R. Rivest, A. Shamir, L. Adleman, *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM, Vol. 21 (2), pp.120-126. 1978.

[16] C. E. Shannon, *A mathematical theory of communication*, Bell System Tech. J. 27 (1948), 379-423 y 623-656.

[17] C. E. Shannon, *Communication theory of secrecy systems*, Bell System Tech. J. 28 (1949), 656-715.

[18] US National Bureau of Standards, *Data Encryption Standard*, Federal Information Processing Standard (FIPS) Publication 46, January 1977.